Vol. 21 No. 4 (New Series No. 82) March 1961

ANALYSIS

THE UNIVERSITY OF MICHIGAN
PERIODICAL READING ROOM

Acting Editor:
P. T. GEACH
with the advice of
A. J. Ayer
R. B. Braithwaite
Herbert Dingle
A. E. Duncan-Jones
C. A. Mace
A. M. MacIver
H. H. Price

CONTENTS

Projection and Paraphrase in Semantics
J. A. FODOR

Will Power
HENRY ALKER

Logical Indeterminacy and Freewill
D. M. MACKAY

Comment on Professor MacKay's Reply
C. J. F. WILLIAMS

Justice and the Common Good
BRIAN M. BARRY

Is 'Good' Ambiguous?
ARNULF ZWEIG

Red and Green All Over Again
PETER REMNANT

Mr. Lehrer on the Constitution of Cans
BRUCE GOLDBERG and HERBERT HEIDELBERGER

THREE SHILLINGS AND SIXPENCE NET

BASIL BLACKWELL · BROAD STREET · OXFORD

ANALYSIS 21.4 MARCH 1961

PROJECTION AND PARAPHRASE IN SEMANTICS[1]

By J. A. FODOR

ONE aspect of the traditional problem of meaning which recent discoveries in structural linguistics may illuminate concerns the ability of speakers of natural languages to compose and interpret indefinitely large numbers of novel sentences at will. There is, of course, no reason to suppose that this nominally linguistic ability to deal appropriately with unfamiliar sentences is in any sense homogeneous. Rather it seems likely that cues of many different kinds are relied upon in deciding what interpretation to place upon token occurrences of sentences. Many of these cues are, in all probability, non-verbal. I have argued elsewhere that to ask what a speaker means *by* a sentence —as distinct from asking what the sentence he utters means—is often to invite reference to aspects of the non-linguistic context of the utterance.[2]

Nevertheless, it seems that some aspects of the ability to understand and produce novel sentences may be characterized in terms of our present knowledge about the systematic relations of sentences in natural languages. To give one example: it is clearly a necessary condition for understanding a new sentence that one should be implicitly capable of giving a grammatical analysis of the sentence. This involves at least the ability to decide correctly upon those substitutions for various components of the sentence which preserve grammaticality. The set of such substitutions, however, partially determines the relation between the sentence in question and other sentences in the language. Thus, understanding

1. The boy went to the store

involves knowing such facts as that

2. The girl went to the store

is a grammatical sentence in English while

3. The gives went to the store

is not.

[1] This paper assumes some acquaintance with the notion of a transformational grammar considered as a device for recursively enumerating the set of grammatical sentences in a natural language. Cf. N. Chomsky, *Syntactic Structures*, Mouton and Co. 'S-Gravenhage, 1957.

[2] J. A. Fodor, '"What Do You Mean?"', *Journal of Philosophy*, 12 July, 1960.

An ability to give grammatical analyses of novel sentences, though clearly *necessary* to understand them, is equally clearly not sufficient. Thus, one might be able to group all the English sentences of a given grammatical type without knowing what any of the sentences so grouped mean. It is not inconceivable that one should know the grammar of a language without knowing the language. For example, among the extra-grammatical abilities involved in understanding sentences, we may count the following: one cannot in general be said to understand a pair of sentences unless one can decide whether or not they are paraphrases of one another.[1]

The ability to decide upon its relations of paraphrase (P-relations) is, I shall assume, one dimension of the understanding of a sentence. To understand a sentence is at least in part to be in agreement with other speakers of the language as to what constitutes a paraphrase of that sentence. The relevance of this point to the problem about how speakers understand *novel* sentences lies in the fact that P-relations, like grammatical relations, seem to be to some extent projectible. One does not have to be trained in the case of each new sentence to recognize its P-relations. Rather, given a knowledge of the P-relations among some proper sub-set of the sentences, it seems possible for speakers to determine the P-relations among the rest of the sentences in a mechanical fashion. In general, the arguments which suggest the recursiveness of the grammatical rules governing natural languages seemt to suggest that the same characteristic belongs to the rules governing P-relatedness. It is this latter system of rules which a theory of semantics must seek to uncover.[2]

At this point we must again be careful to notice the heterogeneity which is characteristic of verbal skills. Not *all* the rules governing P-relatedness are reasonably to be considered linguistic rules. For example, some P-relations are mediated by theories.[3] That

4. An accumulation of ferrous oxide

is P-related to

5. A deposit of rust

is a consequence of chemical theory, not of the conventions of ordinary language. This point need not, however, cause us to despair of the possibility of a linguistic theory of P-relations. Markedly theory-

[1] I use ' paraphrase ' rather than ' synonym ' in part to postpone the usual problems about referential opacity. ' John said that it will rain ' is a paraphrase of ' John said " it will rain " ' though the sentences are not synonyms in that one might be true while the other is false.

[2] For the recursiveness of grammatical rules, cf. H. Putnam, ' Some Issues in The Theory of Grammars ', forthcoming in the *Symposium on Mathematical Methods in Linguistics* published by the American Mathematical Society.

[3] Cf. H. Putnam, ' Dreaming and Language ' (unpublished).

mediated P-relations such as that which obtains between 4 and 5 may in fact *not* be among the relations of a sentence upon which one must be able to decide if one is to be counted as understanding the sentence. Again, though P-relations are often theory-mediated, they are surely not always so. That

6. She admired me

and

7. I was admired by her

are P-related is a consequence not of theoretical connexions but of P-relations involved in the structure of ordinary language. It is P-relations of this kind of which we particularly need an account.

Since it appears that P-relations are decidable for new sentences of arbitrary length (if one has time for computation) it seems plausible to think of the P-relations of long sentences as determined by the P-relations of their components. A theory of P-relations might well take the following form. One assumes as given the P-relations among a finite set of minimally short sentences.[1] (Quine, in a proposal which we shall discuss presently, permits himself only the weak assumption that every such sentence is P-related to itself.) One then requires a set of rules which project these kernels into the set of grammatical sentences. Such a mapping is a transformational grammar of the language. A transformational grammar which preserves P-relatedness in some non-trivial way would, *ipso facto*, constitute a theory of P-relatedness among the sentences of the language. Simplicity of the general theory requires, of course, that wherever possible an intimate relation shall be maintained between the grammatical rules and the semantical rules.

In *Word and Object* Quine makes the following suggestion:

> . . . a concept of synonymy initially significant only for short sentences can be extended to long sentences by analogy as follows. By a *construction*, linguistically speaking, let us understand any fixed way of building a composite expression from arbitrary components of appropriate sort, one or more at a time. . . . Now two sentence-forming constructions may be so related that whenever applied to the same components they yield mutually synonymous results, as long as the results are short enough to be compared for synonymy. In this event it is natural, by extension, to count also mutually synonymous, any results of applying those constructions to identical components however long.[2]

The technique that Quine suggests, if it is feasible, must lead to a theory of synonymy relations among (at least some) sentences, and, *a*

[1] A minimally short or 'kernel' sentence is, tentatively, one from which the deletion of any morpheme renders the sentence ungrammatical. (But what sentences are in fact taken as kernels relative to a given grammar will be decided in part by considerations of the simplicity of the grammar as a whole.)

[2] Quine, W. V., *Word and Object*, The Technology Press of The Massachusetts Institute of Technology, 1960, page 49.

fortiori, to a theory of P-relations among those sentences. There are, however, rather strong objections to attempting to construct a theory in this form.

It should be noticed to begin with that the rules generally employed by linguists for generating a language from kernels do not yield synonymous (or, for that matter, P-related) outputs for applications of alternative transformations (constructions) to a given string. That is, if two transformations are distinct, the sentence which results from applying one of these to a given sentence will not in general be a synonym of the sentence which results from the application of the other transformation to that same sentence. The reason for this is not far to seek. Transformations may be roughly classified as either (a) permutative, or (b) additive.[1] Permutative transformations have the effect of varying the order of the morphemic components of a sentence, and hence obtain between such sentence pairs as

8. The butler did kill the squire

and

9. Did the butler kill the squire?

Additive transformations embed (not necessarily proper) sentence fragments in a given ' matrix ' sentence. Thus they obtain between such pairs as

10. The young man robbed the bank

and

11. The young man whom I knew in Chicago robbed the bank.

Thus, the result of an application of an arbitrary transformation to an arbitrary sentence is either to change the number of morphemes comprising the sentence or to vary their arrangement, or both. A pair of transformations is identical in case they apply to the same sentences and yield the same output.

The point of this is that in general, though not invariably,[2] the performance of either kind of operation, addition or permutation, upon any sentence will change the meaning of that sentence. We cannot usually add morphemes or change their order *ad. lib.* without prejudice to meaning. It is this fact which prompts the belief that the meaning of a sentence (and hence its P-relations) must be in part determined by the set of decisions made at optional choice-points in its transformational generation.[3]

[1] The question whether transformations which delete morphemes may be allowed, and if so under what restrictions, is at present disputed. Upon its solution depend both the recursiveness of the generative rules and the asymmetry of the transformational relationship.

[2] The most important exception is the so-called active-passive transformation, which obtains between such pairs as 6 and 7 above.

[3] I owe this suggestion at least to N. Chomsky, H. Putnam, and P. Ziff.

The extent to which transformation varies meaning may be grasped from the observation that sentences are only rarely synonymous with sentences in their own derivational history. That is, it is generally the case that if S_1 is a sentence transformationally related to S, then S and S_1 are not synonyms.

If, then, transforms of a given sentence are not usually synonymous, is there any way in which P-relations (or synonymy relations) among kernels may be projected in such fashion as to determine the P-relations among constructed sentences? That is, is there any reasonable restriction that can be placed upon putative transformations which will have the effect of systematizing the P-relations in the language? I should like to propose, tentatively, the following as necessary conditions to this end.

We might require of permutative transformations that they be such that if the same ordered set of transformations is applied to two originally P-related kernels, then the resulting output sentences must be P-related. Of additive transformations we might require that if they embed P-related sentence fragments of a given grammatical type in originally P-related matrix sentences, then, again, the output sentences must be P-related.

To put the matter briefly, we can state the difference between Quine's proposal and this one as follows. Quine requires that we find constructions which satisfy the condition that for any sentence S there exist a pair of transformations T_1 and T_2 such that the sentence which results from applying T_2 to S is a synonym of the sentence which results from applying T_1 to S. Our proposal, on the other hand, requires, in effect, that given any set of sentences which are anteceedently known to be P-related, the results of applying identical sets of transformations in the appropriate fashion should be P-related. It is a consequence of both proposals that they permit decisions as to the P-relations of very many sentences, given knowledge of the derivational history of those sentences (and, in the present case, given knowledge of the P-relations among kernels).[1] I wish to claim for my version only that it is a more plausible proposal given what we know of the rules actually employed in transformational descriptions of languages and of the effects of transformation upon meaning.

Magdalen College, Oxford

[1] From the psycho-linguistic point of view it seems reasonable to suppose that the P-relations among kernels are assimilated directly in learning a language. As a point in semantics, however, it is tempting to think of a dictionary as primarily a device for establishing such relations among the basic sentences of a language.

WILL POWER

By HENRY ALKER

ONE aspect of the many-sided problem of free will is concerned with the question whether or not we as individuals do, or as human beings can, have anything such as will power. Three formulations of what this question means will be presented in this paper. I do not propose to answer the question except in so far as that will be necessary to make the meaning of the question, or questions, clear.

To say that an individual's behaviour manifests will power is in some important ways like saying that he had determination or even strength of character. As a good long distance runner nears the finish line fighting off the fatigue that is spreading through his body and struggling to tap every last ounce of reserve energy he can muster, that performance might naturally be described as exhibiting will power or strong determination. If the runner's behaviour was, as we might expect, the result of long training, conscientious self-discipline and part of his dedication to a healthy and vigorous way of life, we might even say that it evinced strength of character. From these simple points we may surmise that 'will power' can be used as a term of appraisal. A *bad* long distance runner as he was completing the last lap might succumb to the fatigue permeating his limbs, and find himself unprovided with the strength to carry on energetically to the finish line. If this failure resulted from generally sloppy training and, say, an occasion the night before when, just for the hell of it, the runner knocked back several rounds of hard liquor, then it could be interpreted in terms of a lack of will power, or, perhaps, as a manifestation of weak character. The interesting point now arises whether or not the ascription of will power to a person need always be a matter of appraising his actions in some way. Does there exist a purely descriptive use of that concept?

Certainly in the case of determination we could answer this question in the affirmative. Consider, for instance, the example of a man who with quiet determination counts the number of lamp posts in London. He gets up an hour earlier than he needs to every morning just so that he can take a different route to work and check the number of lamp posts located on the way. He arrives home an hour later than he needs to for the same reason. Furthermore he double-checks each street on some second occasion to make sure that he has not miscounted. No special anxiety arises if circumstances make it physically impossible to continue his counting on some occasion. He is not obsessed with his task but just determined to complete it. So far the statement that he goes about his self-initiated undertaking with determination could serve to qualify him as a good lamp post counter. But suppose that a very reliable

survey on the same subject had been made the year before. And suppose that our man was perfectly well acquainted with the existence and reliability of this report. Consequently in this case the determination does not count towards making our man a good lamp post counter. We could fill in the example with features that made it clear, also, that he did not intend sometime in the future to count lamp posts in a city where no such survey was available. So the determination does not count towards making him a good potential lamp post counter. When the object of his determination appears in no way praiseworthy or blamable, the determined way he goes about accomplishing it need not make for any difference in the way we appraise his behaviour.

The justification for dwelling on such a silly example is that it helps to highlight an important and closely related question about the concept of will power. Can we say of it too that it may be used in such a way as to describe activity without in any way carrying the implication that the actions so produced are worthy of praise or blame? We can answer this question by adding a few details to the example sketched above. Suppose in that example the effort made by our non-heroic man includes more than the minor exertion required to get out of bed in the morning an hour early and refrain from going home for an hour after work at the office is done. Let us say that he often is quite hungry after leaving his job but nonetheless forgoes this pleasure to get his lamp posts counted. Occasions arise when the desire to eat manifests itself quite strongly. Our non-hero lets his eyes linger over the menus in the windows of several restaurants he passes, salivates abundantly when aromas from these places reach his nose, and so on. We could say of this man, as he repeatedly and vigorously turns his head away from the restaurant windows and rechecks the number of street lamps to make sure that he has not miscounted due to the momentary break in his concentration, that here is an example of will power. The ridiculous air of the example suggests that no imperative import attaches to the description of it by means of the term ' will power '.

Now we are in a position to understand how will power may be manifested in the behaviour of human beings, and on the way to discovering whether or not it is present in a particular person. By elaborating our analysis we could show that the phrase ' will power ' may be applied fairly clearly to a certain range of cases. Its application resembles in some ways and differs in others (which I have not mentioned) from the use of phrases like ' determination ' or ' decision of character '. The briefly sketched examples I have given help towards providing a first possible sense to the assertion that will power exists. And in so far as the question ' Does will power exist? ' is a request for some typical or boundary line illustrations of how the concept is used, they simultaneously work towards answering that question.

Any deep-seated philosophical perplexity about the matter, however, probably still remains unresolved. The question could arise following such an analysis whether or not the concept as it actually was used was correct from a scientific point of view. For, as was particularly evident in the first example, talk about behaviour manifesting will power is also talk about the manner in which that behaviour has been produced. This second variety of doubt about the existence of will power might be developed as follows. In effect, the analysis of will power given at the beginning of this paper constitutes a paradigm-case argument for the existence of that entity. This kind of argument cannot establish the existence of will power in cases where users of the concept know perfectly well what they mean by the term. We are dealing with a theory-embodying concept. It incorporates certain assumptions about factual matters which the scientist has a right to declare empirically impossible or empirically false. The theory presupposed by the particular concept in question resembles an out-of-date faculty psychology in which Reason, Feeling and Will are hypostatized as separate agencies or faculties governing our behaviour. That theory might very well be falsified by research carried out under the aegis of the modern experimental approach to personality. Probably it would also be shown by the scientist or the philosopher of science to possess a pseudo-explanatory persuasiveness. That the concept of will power has an application does not at all guarantee that it should be applied. To move from the assertion that will power exists in the first sense specified above to the assertion in the second sense we are now considering, is to move from the elucidation of ordinary usage to the elaboration of common sense.

As indicated above it is not the intention of this paper to answer the questions about the existence of will power raised herein. Nonetheless the above argument will be developed a little further in order to explicate more clearly what kind of question about the existence of will power we are examining. A reply to the above denial of will power could take several directions. It might contend that the notion of will power did not contain any kind of empirical theory or at least not the particular theory we have suggested that it contains. It might also argue that the theory embodied in the concept is not at odds with what behavioural scientists (physiologists, psychologists, or political sociologists) know about will power, although admittedly it does not possess the precision present in the analysis of will power from those points of view. Both the philosopher and the scientist would be needed to sort out and assess this argument about technical and non-technical concepts.

We are now ready to raise a third question about the existence of will power. This thesis is concerned with the implications, for the user of moral language, of the second thesis sketched above. It narrows the question to the use of the concept in contexts of moral appraisal.

Assuming that some factual content is presupposed on occasions when we praise a person for the abundance of his will power or blame him for his lack thereof, should we go on to say that such an appraisal is out of order until the scientific evidence is consulted? Might we not find, this query continues, that praising a person for some special power or force he cannot have (e.g. will power) renders that appraisal meaningless? It would be absurd to evaluate a person on the basis of whether or not he could flap his hands and fly or was born in such and such an astrological period. In a similar way is not the evaluation of a person in terms of his will power equally inappropriate? An interesting point about this argument is that within the general context of the problem about free will we have arrived at a point where the consultation of scientific opinion is required. For this question depends, in part, on what the scientist might say about the second thesis about the existence of will power we have identified above. It depends *only* in part, however, on such a consideration, because the user of moral language could say, for instance, that whatever the scientist did say about the contrast between cases where one surrendered to the temptations of the flesh and cases where will power enabled one to resist iniquity, he would still continue to insist that the one action was good and the other bad. And if people consulted a scientific attack on the existence of will power, agreed with it, and still conscientiously maintained their previous position, then their predicament could still give rise to genuine questions.

In conclusion let us note one relation between our problem and the philosophical puzzle about free will. We could derive from the three ways of asserting the existence of will power mentioned here three analogous ways of asserting or denying the existence of acts of the will.

Magdalen College, Oxford

LOGICAL INDETERMINACY AND FREEWILL

By D. M. MacKay

MR. C. J. F. WILLIAMS (Analysis 21.1 (October 1960), pp. 12–13) is right in pointing out some analogy between the logic of (P_1) ' Provided A does not get to hear of this, he is certain to do X ' and that of (P_2) 'A is mistaken in thinking *p*'. In each case, A is logically precluded from believing as ' true now ' what is asserted secretly by the observer. It is easy in fact to construct a large number of roughly analogous assertions, by making some aspects of the communication process itself a subject of the assertion.

Mr. Williams thinks that the demonstration of this analogy disposes of my argument (*Mind* lxix (January 1960), pp. 31–40) that what the isolated observers may validly believe about A's future action cannot claim *ipso facto* to represent ' the true state of affairs ', and that A's relation to what they believe cannot properly be described as ' ignorance '. " If the proposition that A will do X is true for anybody it is true for everybody, even for those who cannot express it " (Williams, p. 13).

Now (oddly enough) I am so far from disputing this last dictum, that my argument can actually be summed up by its converse: if a proposition is demonstrably *un*true for anybody, it *cannot* claim to be true for everybody. My only use of the expression ' true *for the onlookers* ' (my p. 35) was in order to *weaken* the sense in which I would allow their belief to be called ' true '. To express this weakened status I thereupon substituted the term ' valid ', and I did not in fact use the expression ' true for A' on which Mr. Williams (p. 13) says my arguments rest.

To speak of a belief as ' *valid* for A' but not ' valid for others ' is perfectly meaningful, as the daily usage of relativity theory demonstrates (see my p. 37). It implies (*inter alia*) that A's expectations will not be let down, whereas those of others, if organized by the same proposition (not just the same *sentence*, as Mr. Williams would have it), would be falsified.

But of course, to return to Mr. Williams' dictum, my objection to calling 'A will do X ' *true*, is not that A ' cannot express ' it, but that (*ex hypothesi*) A would be *wrong to believe it*. This is something far deeper than the need to reformulate ' John kicked James ' when it is John who speaks. To quote from my paper (p. 39) " it is *what is asserted* by the onlookers, and not just its conventional *form*, that is logically unacceptable to the agent ".

What then of the analogy with P_2: 'A is mistaken in thinking *p* '? Let us follow through. We believe P_2 to be ' true ', and urge it cogently on A. Given our data, A is convinced. He agrees to renounce *p*. If

asked about P_2, he will say ' Oh yes, I *was* mistaken. P_2 was valid, so I can now accept it (if labelled for time) as true of me-then (*N.B. not* true-then of me) '. There the matter rests. We have won our point.

Now let us try the same game with our P_1. This time, no matter how cogently we argue, A not only refuses to be convinced but confidently defies us to predict his action. We try it in the form ' Provided A does not get to hear of this, he is certain to do X '—but of course A dismisses this as empty and out of date. The argument reaches—and can reach—no resting-place.

Why? For the simple reason that the time-label we wish to attach to P_1 names a time *still in the future*, whereas in the case of P_2 the time named is in the past. What gives P_2 its stable truth-value is the fact that it, and its grounds, can be discussed with A *only in the past tense*. What gives P_1 its logical indeterminacy for A is that the *data* on which it is based cannot be accepted by A without self-invalidation, for they are changed by his acceptance of them—they come, if you like, from his present rather than his past. They are always logically one jump ahead of his contemplation of them.

To point the contrast still further, we may note that the proper analogy of P_2 is not P_1, but something like (P_3): 'A is in a state which will entail X if A does not hear of it '. This proposition can indeed be urged on A, with (we suppose) adequate support from past cinephotographs of his brain and the like. And I see no objection to A's assenting to it any more than to P_2. But as far as his choice (still in the future) is concerned, he will of course again dismiss the exercise as ' out of date ', since he is no longer in the state described in the data offered. With regard to his future action, he would still be *wrong* (and not just unable) to believe any proposition of the form P_1, let alone of the form 'A is certain to do X '. And in agreement with Mr. Williams' own dictum, it would seem that this releases such a proposition from being called ' true '. To call it ' valid ' for the isolated observers, and ' logically indeterminate ' for the agent—until after the event—offers I think a useful way of reserving ' true ' for those propositions that do fulfil the dictum.

University College of North Staffordshire

COMMENT ON PROFESSOR MACKAY'S REPLY

By C. J. F. WILLIAMS

PROFESSOR MACKAY claims that the analogy which I pointed out between the sentences which he has labelled 'P_1' and 'P_2' breaks down at two points which are relevant to our argument. First, he claims that, whereas A's position in regard to P_2 is merely that he cannot express the proposition it states, in the case of P_1 he would be *wrong to believe* (his italics) the proposition stated by it. (My own description of A's position in regard to P_2 was incomplete. I said (p. 12) that P_2 'expresses a proposition which it is impossible for A himself to express': I should have added 'by means of a sentence containing a part of the verb "to be mistaken" in the present tense'. For 'I *was* mistaken in thinking *p*' is a sentence which, given certain circumstances of time and persons, might well be used by A to express the same proposition as that expressed by P_2.) Now MacKay is quite right in saying that A's position in regard to P_2 is merely as I have described it, and that we should not say 'A is *wrong to believe* that he is mistaken in thinking *p*'; for this implies 'A believes that he is mistaken in thinking *p*', which is self-contradictory. But neither can we say, as MacKay wishes, 'A is wrong to believe that, provided he does not get to hear of this, he is certain to do X'. (I am here assuming that MacKay's vacillation between P_1 and 'A will do X' is unintentional. It is 'A will do X' which he actually says (p. 82) A would be wrong to believe, but I take it that he wishes the same to be said of the proposition that is relevant to the discussion, namely, that expressed by P_1.) The only reason for thinking that A is in a peculiar position with regard to P_1 is that it would be self-contradictory to say that A believed it, and *a fortiori* to say that he was wrong to believe it.

Secondly, MacKay describes a situation in which A is brought to see the truth of P_2. Although we cannot convince A that he *is* mistaken in believing *p*, we can convince him that he *was* so mistaken. But he does not think we shall succeed if we 'try the same game' with P_1. Why not? Although we cannot convince A that he *is* certain to do A provided he *doesn't* get to hear this, we can convince him that he *would have been* certain to do X *had he not* got to hear. The parallel is exact. MacKay says that P_2 has 'a stable truth-value' because 'it, and its grounds, can be discussed with A *only in the past tense*'. If, therefore, we can discuss with A whether he would have been certain to do X provided that he had not got to hear of the prediction, this is enough, in MacKay's words, to give P_1 'a stable truth-value'.

The only thing that prevents A's assenting to P_1 in its original form

is that its being expressed in this form implies that A has not heard about the prediction. If it were true that A *had* got to hear of this, we could not say ' Provided A does not get to hear of this, he is certain to do X ', but would have to recast the hypothetical in subjunctive form. A is prevented from saying something which implies ' I have not heard of q ' for reasons analogous with those that prevent him from saying ' I am mistaken in thinking p '.

I suspect that MacKay is confusing the logical impossibility of A's saying ' Provided I do not get to hear of this I am certain to do X ', with the causal impossibility (inside MacKay's original story) of A's hearing about the prediction without the data for it being upset. If the fanciful account of psychological determinism which MacKay has given us were correct, A would indeed be ' wrong to believe ' that it was certain that he would do X, because, as a matter of fact, the only data on which the reasoning could be based included his ignorance of the prediction. That the data are of this sort, however, is by no means a necessary truth. They might well have consisted solely in facts about heavenly bodies, had MacKay chosen to tell a different story. In that case, supposing that A was a sufficiently competent astronomer, he might have been perfectly right to be certain that he would do X, without in any way violating the laws of logic.

A cannot say (P_1) ' Provided I do not get to hear of this I am certain to do X '. He can, however, say ' Provided I had not got to hear of this I would have been certain to do X '. This is enough to establish the truth of the contradictory of the remark from the original paper (p. 39) which MacKay quotes in his reply (p. 82): it is *not* ' *what is asserted* by the onlookers' but only ' its conventional *form*, that is logically unacceptable to the agent '.

Balliol College, Oxford

JUSTICE AND THE COMMON GOOD

By BRIAN M. BARRY

I

Social Principles and the Democratic State, by S. I. Benn and R. S. Peters (George Allen & Unwin, London, 1959) is far more than a textbook; for the authors' object is not merely to say, ' Justice means this, equality means that, freedom means the other. Where they conflict you take your pick '. Their thesis is the daring one that all political arguments fit into a single pattern and that this pattern is identical with morality.

The authors reject any sociological definition of ' moral rule '. According to them, a rule is moral if and only if it is:

> critically accepted by the individual in the light of certain criteria. The criteria can be summarized by saying that a rule should be considered in the light of the needs and interests of people likely to be affected by it with no partiality towards the claims of any of those whose needs and interests are at stake. (p. 56)
>
> People have interests and needs which they put forward as claims. The criteria of impersonality and respect for persons are satisfied when claims are assessed on relevant grounds, and privileges excluded as a basis for allowing a claim. (p. 51)

I think the *inconvenience* of this definition is fairly clear: it follows from it, for example, that one can never talk about the moral rules of a group unless one has first ascertained whether everyone who acts on the rules and expects others to act on them has accepted them after an impartial consideration of their effects. The theoretical disadvantages of making stipulative restrictions on the expression ' moral rules ' are very similar to the disadvantages (which the authors recognise) of calling only good laws ' laws '.

My main objection, however, is that no single scheme can be sufficient to cover all the arguments which even in liberal-democratic communities would ordinarily be thought ' moral '. If Benn and Peters do succeed in interpreting their schema so as to fit any argument presented into it, this is at a cost of both misrepresenting most arguments and making the schema vacuous.

II

This must be so, I suggest, because when we are dealing with interests there are two conflicting principles at work: aggregative and distributive. They are both, it seems to me, independently operative in most men's minds; and where they give conflicting answers there is no higher principle to which the conflict can be referred.

Suppose, for example, that one can see no reason in terms of desert or need why science teachers should be paid more than others; but that one also believes that unless all the money available for raising pay goes to them the standard of science teaching will decline with grave results. Or again, suppose that one believes on the one hand that dropping two atomic bombs on Japan will cause less suffering than continuing the war with 'conventional' weapons, but that it is unjust to use weapons which rely on devastating civilian populations rather than military targets. In both cases we have a conflict between an aggregative result and a distributive principle. If these were all the relevant considerations (which in fact they are not) then how one decided would depend on how highly one ranked the two kinds of principle (moral philosophers can be divided according to which side they exaggerate at the expense of the other).

The point I wish to make is that 'impartiality' is no help here in providing a schema for decision. If it means 'everyone to count for one and nobody for more than one', this is satisfied by the result of aggregation, which may still be morally unacceptable to many people. If on the other hand one is to say, 'The arts teachers aren't given *enough* consideration by straight aggregation; they must be given some more', my objection is that the formula is now useless as a guide to decision. It merely gives us a thoroughly misleading way of justifying *ex post facto* whatever decision we may in fact reach.

III

There is still, however, one possible escape route for the authors' formula. This lies in emphasising the bit about 'relevant grounds for treating people differently'. But this manoeuvre fails in one of two ways, depending on the construction of 'relevant'. If you say that what subject someone teaches is not a relevant criterion in determining his pay, you certainly get the answer that scientists ought not to be paid more. But this is too good a demonstration, for according to the authors, anyone who admits this and still says that science teachers ought to be paid more is not arguing morally; yet such a position seems to me perfectly reasonable. Some one who takes this view has chosen in favour of the education of children and against professional equity—surely a perfectly reputable thing to do. If on the other hand you say that what subject someone teaches is made relevant simply by the fact that science teachers are in shorter supply than others and paying them more will keep up educational standards, you are right back at a simple aggregationist position.

In fact the authors give some support to each interpretation. Thus, on p. 112, they say '. . . it is up to whoever should make distinctions to

justify the criteria in terms . . . ultimately of a balance of advantage to all concerned '. Phrases such as ' beneficent consequences ' and ' beneficial results ' occur in a similar context on pp. 169 and 170. This sounds like a straight aggregationist position: you show that something is just, or for the common good, by showing that it makes relevant distinctions, and these distinctions are relevant if they provide ' a balance of advantage to all concerned '. This form of argument is certainly used by the authors and I shall discuss an example in section IV.

Against this, they say on pp. 272–3:

> Two politicians may each say, with perfect sincerity, that he is seeking the public interest, or the common good, though one proposes to expropriate private capital and the other to defend it for the death. Does one of them have to be wrong. . . ? Is the disagreement about fact at all? It may be . . . but the probability is that this is less important than a disagreement on moral principle. One holds private capital to be an immoral thing in itself, the other that it represents the legitimate fruits of thrift, industry, and other economic virtues. . . . What then have the two politicians in common, that enables them to appeal, with equal sincerity to ' the common good '? . . . (Each) is saying, in effect, that having considered the claims of *all* sections in a spirit of impartiality, the balance of advantage lies in the course he recommends.

Here it is quite clear that ' the balance of advantage ' is not something obtainable by aggregation; it is simply a repetition of the procedural point that one must have considered all claims impartially. To say on this analysis that relevant distinctions are those based on a balance of advantage is to add nothing. ' We have to *decide* what is relevant '. (p. 113).

My object in pointing out this inconsistency is not to score a cheap debating point but to substantiate my view that it is impossible to fit into one theory questions of distributive principle (the sort of thing the two politicians in the argument are disagreeing about) and aggregative ones (which phrases such as ' balance of advantage ' would naturally be thought to refer to):

IV

In this section I shall apply the above analysis to two of the discussions of particular questions in the book to illustrate how the authors' insistence on trying to show all concepts as aspects of a single criterion leads them to distort characteristic forms of argument. In the first example, we see a straight question of distribution obscured by reference to an aggregative concept. On p. 272, the authors say:

> . . . the government would have resented being told in 1957 that decontrolling rents was not for the common good. But the government clearly had to choose between the interests of the landlords and the interest of the tenants. Whether or not it chose rightly, it did little good to the tenants,

Now if it is correct that the issue is one where one side's losses are the other side's gains, I would suggest that the ' common good ' is out of place. *Of course* the government would resent being told its action was *not* for the common good; but neither would it justify the measure by saying it was *for* the common good. The concept is out of place. If the government is willing to admit that it is simply transferring money from one set of pockets to another, it will say, for example, that control was unfair between owners of different forms of property, or between owner-occupiers and tenants, or that it was unjust between landlords and tenants; i.e., it would support a distributive change by distributive arguments. (More likely, of course, it would deny that only one side would be benefited, pointing to the benefits of a free market in producing a rational allocation of resources, and then it *could* talk about the common good; but I follow the authors' assumption for the sake of the argument.)

What I am trying to show is that it is not an *accident* that we have different concepts; they really do have different jobs. We have one set which point out various distributive comparisons, such as justice, fairness, equity, equality (that these all differ is not hard to show but it would take me out of my way here. One example: a lottery is fair if honestly run, but a lottery which distributed prizes *justly*, i.e., according to desert or need, would no longer be fair). And we have another set which point out the results of various methods of aggregation, such as ' public interest ', ' common good ' and ' general welfare ' (' good ' and ' interest ' for example require one to include different ways in which people are affected). To say, as Benn and Peters do, that ' to seek the common good ' means ' to try to act justly ', is to make nonsense of the subtle and complex way in which we go about criticizing political programmes, in the pursuit of a tidy but barren theory.

My second example shows the same error operating in reverse. In chapters 4 and 5, the authors have an excellent study of the grounds on which one might justify various claims to income. Unfortunately, however, they are again hampered by their theoretical apparatus, for all claims, according to their general theory, must be established by being shown to be just. This works excellently for claims based on personal desert and need; but it does not work at all for arguments for property based on the advantage of having a group of politically independent or cultured citizens (pp. 167–70):

> If it were true that fortunes based on inheritable property were indispensable for an elite of this sort, and if such an elite were really so valuable, the property system would be justified by its beneficient consequences. (p. 169)

Now, this is fine; and the situation should in my view be summed up by saying that in this case justice would have to be qualified by utilitarian considerations.

But this course is not open to Benn and Peters. They have to say that the general advantage of property makes the amount of money your father had a relevant and therefore just ground for differences in income (pp. 169–70). This seems to me highly misleading. Although Hume used the expression ' rules of justice ' to cover precisely such things as property rules, ' justice ' is nowadays analytically tied to ' desert ' and ' need ', so that one could quite properly say that some of what Hume called ' rules of justice ' were unjust. Again, we see how the attempt to reduce all arguments to one pattern forces Benn and Peters to assimilate quite different kinds of arguments to one another.

University of Birmingham

IS 'GOOD' AMBIGUOUS?

By ARNULF ZWEIG

ARE there specifically ethical words? Many philosophers seem to think so, or at least to hold that words like 'good', 'right', and 'ought', when used in certain contexts, designate specifically moral qualities, or have a specifically moral meaning or use. Other philosophers deny that there are distinctively ethical words, or uses of words. One way the question can be stated is by asking whether words like 'good' are ambiguous because we speak of good behaviour and good apple pie, giving one sort of reason to back up the claim that a pie is good and other sorts of reasons on behalf of moral appraisals. Or are we confronted merely with different kinds of valuables, the value term itself being unambiguous? C. H. Whiteley has argued recently that 'good' is not ambiguous.

> But words like 'good', 'right', 'ought', which are used to evaluate and to recommend in the discussion of moral issues, are also used, *without change of meaning*, in the discussion of other matters. Conduct can be good, as meat can be good. The fact that I can approve of conduct on moral grounds . . . no more makes 'good' an ethical word than the fact that I can approve of meat on dietetic or gastronomic grounds makes it a dietetic or gastronomic word.[1]

The argument will appear cogent to some philosophers, especially if one thinks it useful to distinguish what a word means from the criteria on the basis of which we apply it. But even if one grants that distinction, one may hesitate to accept the suggestion that words like 'good' do not differ in meaning when applied in contexts of moral appraisal from what they mean in other contexts. My point is not that they do, but rather that we have no clear way of showing that 'good' does or that it does not change its meaning when applied to different classes of objects or events. If one wants to maintain that there are distinctively moral words or moral uses of words, one will interpret contexts of moral appraisal (however identified) as not simply a sub-set of the class to which the putative moral words may be applied, but as constituting a class which differs in *kind* from other classes.[2] In other words, one *may* regard 'good' as a moral word in discussing someone's behaviour and (what Whiteley denies) as a gastronomic word in discussing his food. We are not compelled to regard the word as univocal; whether we do or not will depend obviously on our prior convictions as to what 'good' means.

[1] "On Defining 'Moral'", ANALYSIS, 20.6 (June 1960), p. 144.

[2] This point has been defended in some detail by my colleague, Professor R. J. Richman. See his article, 'Ambiguity and Intuition', *Mind*, lxviii (1959), p. 269

Realizing this should make one cautious in claiming that a word like ' good ' is or is not being used " in a different sense " on one or another occasion of its use. Where some see species of goodness, others find clear-cut ambiguity. Everything seems to depend on *how* different the goodness of a dinner and the goodness of a charitable act strike us as being. Without some specification of what is to count as evidence for a difference in meaning, the claim that ' good ' is not ambiguous has no more warrant than the claim that it is.

University of Oregon

Dr. Whiteley writes:

> Mr. Zweig's point is sound, in that ' the same meaning ' is an imprecise expression. Since ' good ' used to commend food and ' good ' used to commend character refer to different grounds of commendation, one could say that they differ in meaning; and then there would be as many different meanings of ' good ' as there are kinds of commendable objects. By parity of reasoning, if two people whose moral standards are quite different use ' good ' to commend character, one might say that they mean different things by ' good '; and there would be as many different meanings of ' good ' as there are sets of moral standards. On the other hand, the commending function is common to all uses of ' good '; it is this common element of meaning which makes ' good ' one word instead of a cluster of homonyms, and enables us to understand it in cases where we do not know what the grounds of commendation may be. And this is sufficient reason for saying that it has the same meaning in moral and non-moral contexts.

RED AND GREEN ALL OVER AGAIN

By PETER REMNANT

> '. . . the shortness that is in us declines ever to become or be tall; nor will any other quality, while still remaining what it was, at the same time become or be the opposite quality; in such a situation it either withdraws or ceases to exist.' Plato: *Phaedo*.

'CAN a thing be red and green all over?' In an issue of ANALYSIS of some years ago (December, 1954) Professor A. D. Ritchie remarks that this 'should mean "Can we, in practice, make it red and green all over at one and the same time by any operations with pigments, coloured lights, etc.?"', and objects to the atmosphere of *a-priorism* in which this question is ordinarily discussed.

Let us refer to the proposition that no thing can be both red and green all over as 'P'. I shall maintain that P is certainly true and that it is in some sense analytic, but I shall also attempt to show that discussion of P is complicated by the presence of certain interesting empirical factors.

There are familiar difficulties about classifying P as analytic—certainly it is not a proposition in which 'the predicate B belongs to the subject A, as something which is (covertly) contained in this concept A' (Kant). Nor, in some more liberal sense of 'analytic', does it appear to be a proposition whose truth can be demonstrated simply by unpacking the contents of the concepts which it involves. The principal difficulty here is that words like 'red' and 'green' cannot be unpacked in the way that words like 'uncle' can; since they are introduced ostensively they cannot be resolved into conjunctions of simpler concepts comparable with 'male sibling of a parent'.

On the other hand there is a strong presumption against P being a straightforward empirical generalization. Even with the best confirmed empirical generalizations the possibility of counter-instances can still be made intelligible—I can easily imagine flying pigs. But what am I to conjure up if asked to imagine a counter-instance of P?

Let us look more closely at P. An object cannot be both red and green all over—it can be red and hot, or green and aromatic, but it cannot be hot and cold all over, or at once both bland and aromatic. Some pairs of properties appear to be naturally incompatible with each other—to repel or neutralize each other.

Consider for a moment the kind of repulsion which exists between certain pairs of colours. That red and green are thought of as incompatible is not surprising—when combined by mixing pigments or by superimposing filters they are complementaries which cancel each other

out. Mix yellow and blue, or green and blue, in one of these ways and the result is a new shade of colour more or less intermediate between the original pair—green in the first case and blue-green in the second. But mix red and green in the same way and, depending upon the balance between the two, the result will be a greyish red or a greyish green or just a pure grey with nothing of either red or green in it.

I shall speak of some colour mixing as ' constructive ' and of some as ' destructive '—of some pairs of colours as compatible and of others as incompatible. Some colours when combined produce a new shade of colour which is intermediate between the original two and resembles both its progenitors. Some, like red and green, tend to cancel each other out when mixed and to produce not a new shade of colour but a shade of grey from which all colour has disappeared.

One is tempted to say that the reason why a surface cannot be both red and green all over is that these two colours are incompatible—their mixture is a destructive one. And this is clearly an empirical matter.

But suppose further research of the kind called for by Professor Ritchie yielded a method of constructively combining red and green; would this then falsify P by enabling us to produce objects which were red and green all over? Not at all—blue and yellow are compatible, but the result of painting something with a mixture of blue and yellow pigments is a green surface, not a surface which is both blue and yellow all over. Nor, by painting something with a mixture of blue and green, do we obtain a surface which is both blue and green at one and the same time—the result is a single shade of colour, blue-green. Thus we may conclude that, although constructive mixture of blue and green pigments is physically possible whereas constructive mixture of red and green pigments is not, a surface which is both blue and green is no more possible than is one which is both red and green. The fact is that red and green can be combined constructively by projecting red light and green light onto the same area of a screen. But the result is not, of course, a surface which is both red and green—it is one which is yellow.

Is the impossibility of producing a surface with two colours all over a physical impossibility? Again the answer is no—it is a consequence of the convention which governs the manner in which colour-descriptions are employed. That this is a matter of convention can be seen by recognizing the possibility of employing a quite different convention—one in accordance with which surfaces which were red, or green, or violet, would be said to have one colour, whereas surfaces which were yellow, or blue, for example, would be said to be both red and green, or violet and green, respectively, and so on.

In terms of such a convention a surface would be said to possess two colours when its coloration was the result of the *constructive* combination of two primary colours. Such a convention might be a trifle

inconvenient for everyday use, but would approximate quite closely to the method used by scientists to specify shades of colour.

This amounts to saying that the impossibility of a surface being both red and green is essentially a logical impossibility, although there are empirical impossibilities about colour-mixing closely adjacent to it, and that this logical impossibility arises from the convention according to which any surface which is homogeneously coloured is said to possess only one colour even in cases in which that colour must be described by the use of two colour words, such as blue-green. And to say that a principle is true by convention, or true in virtue of the logical impossibility of its negation, is to say that it is analytically true, at least upon a reasonably liberal interpretation of that ill-defined concept.

But suppose someone now says: 'But the possibility or impossibility of red-green surfaces and the conventions for talking about them have nothing to do with the case; what I am talking about is a surface which is just red—nothing but red—and about the possibility or impossibility of that surface also being green.' Well, I have contemplated the possibility of producing a surface which is in a sense both red and green and the possibility of establishing a convention in terms of which such a surface would be said to be both red and green. But if I am now being asked why a surface which is just red—not in any sense red and green also—could not be green as well, I appear to be being invited to contemplate the possibility of an explicit self-contradiction.

University of British Columbia

MR. LEHRER ON THE CONSTITUTION OF CANS

By BRUCE GOLDBERG *and* HERBERT HEIDELBERGER

MR. KEITH LEHRER, in a recent issue of this journal,[1] attempts to demonstrate that if ' I can ' is analyzed as ' I shall, if I choose ' then ' I shall, if I choose ' does not express a causal relationship between the antecedent and the consequent. We shall not, in this paper, take issue with Lehrer's conclusion, viz., that ' I shall, if I choose ' is not causal; what we shall try to show is that his argument does not establish that conclusion. We shall do this by showing that in his demonstration Lehrer makes inferences which are forbidden by his initial supposition.

The supposition is that ' I can, if I choose ' is not causal. (This, as Lehrer points out, was proved by Austin in ' Ifs and Cans '.[2]) Lehrer's argument, briefly, is this. If ' I can ' is equivalent to ' I shall, if I choose ', then this last expression may be substituted for ' I can ' in ' I can, if I choose '. We thus get ' I shall, if I choose, if I choose ', which is transformed into simply ' I shall, if I choose '. The argument thus shows that ' I shall, if I choose ' is equivalent to ' I can, if I choose ', if ' I can ' is analyzed as ' I shall, if I choose '. But now since ' I can, if I choose ' is not causal, ' I shall, if I choose ' cannot be causal either.

However, the deduction is valid only if it is assumed that the ' if '-clause in ' I can, if I choose ' is a conditional clause. But this is an assumption to which Mr. Lehrer is not entitled, if he accepts Austin's argument —as in fact he does. Austin shows that ' I can, if I choose ' is not a causal conditional, since it is not a conditional at all; for, he argues, if it were, ' If I cannot, I do not choose to ' would be inferable from it, and plainly it is not. But if ' I can, if I choose ' is not a conditional, then none of the steps in Mr. Lehrer's argument can be taken, since they are founded upon the rules of inference sanctioned by the treatment of ' I can, if I choose ' as a conditional.

What Mr. Lehrer has established is that if one construes both ' I can, if I choose ' and ' I shall, if I choose ' as causal, where the latter constitutes an analysis of ' I can ', then one is committed to their equivalence. But if one takes only ' I shall, if I choose ' as causal, then this view is not at all affected by Mr. Lehrer's argument.

Princeton University

[1] ANALYSIS, 20.6 (June 1960), pp. 122–124.
[2] In *Proceedings of The British Academy*, Vol. XLII, 1956, pp. 109–132.